



Draft Genome Sequence of the Yeast *Pachysolen tannophilus* CBS 4044/NRRL Y-2460

Liu, Xiaoying ; Kaas, Rolf Sommer; Jensen, Peter Ruhdal; Workman, Mhairi

Published in:
Eukaryotic Cell (Online Edition)

Link to article, DOI:
[10.1128/EC.00114-12](https://doi.org/10.1128/EC.00114-12)

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Liu, X., Kaas, R. S., Jensen, P. R., & Workman, M. (2012). Draft Genome Sequence of the Yeast *Pachysolen tannophilus* CBS 4044/NRRL Y-2460. *Eukaryotic Cell (Online Edition)*, 11(6), 827.
<https://doi.org/10.1128/EC.00114-12>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Draft Genome Sequence of the Yeast *Pachysolen tannophilus* CBS 4044/NRRL Y-2460

Xiaoying Liu, Rolf Sommer Kaas, Peter Ruhdal Jensen and
Mhairi Workman
Eukaryotic Cell 2012, 11(6):827. DOI: 10.1128/EC.00114-12.

Updated information and services can be found at:
<http://ec.asm.org/content/11/6/827>

REFERENCES

These include:

This article cites 8 articles, 2 of which can be accessed free at:
<http://ec.asm.org/content/11/6/827#ref-list-1>

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new
articles cite this article), [more»](#)

Information about commercial reprint orders: <http://journals.asm.org/site/misc/reprints.xhtml>
To subscribe to to another ASM Journal go to: <http://journals.asm.org/site/subscriptions/>

Draft Genome Sequence of the Yeast *Pachysolen tannophilus* CBS 4044/NRRL Y-2460

Xiaoying Liu, Rolf Sommer Kaas, Peter Ruhdal Jensen, and Mhairi Workman*

Center for Systems Microbiology, Department of Systems Biology, Matematiktorvet, Technical University of Denmark, Kongens Lyngby, Denmark

A draft genome sequence of the yeast *Pachysolen tannophilus* CBS 4044/NRRL Y-2460 is presented. The organism has the potential to be developed as a cell factory for biorefineries due to its ability to utilize waste feedstocks. The sequenced genome size was 12,238,196 bp, consisting of 34 scaffolds. A total of 4,463 genes from 5,346 predicted open reading frames were annotated with function.

The yeast *Pachysolen tannophilus* was first isolated from wood extracts used in leather tanning (2) and has gained interest due to its ability to utilize D-xylose (1, 7, 8) for fuel ethanol production (6, 8, 10). It can ferment the common sugars (glucose, mannose, and galactose) occurring in hemicellulose hydrolysate mixtures (8) and can produce ethanol from glycerol (4). The whole-genome sequencing of *P. tannophilus* was performed to provide genetic information as a necessary step toward engineering of the strain and its potential development as an industrial cell factory.

Sequencing was performed by a whole-genome shotgun strategy with an Illumina genome analyzer (Beijing Genomics Institute, Shenzhen, China). The raw data of short reads were assembled into 279 contigs which were ordered into 34 scaffolds (>2 kb) with an N₅₀ size of 1.1 Mb by using the SOAPdenovo package (3). Augustus software, version 2.5 (9), trained for predicting genes in *Debaryomyces hansenii*, was utilized to identify protein-coding genes in the genome, and the putative amino acid sequences were used for subsequent gene function annotation analysis. The functional annotation was accomplished by BLASTP analysis (E-value < 1 × 10⁻⁵) of protein sequences in the databases (COG and KEGG), and the best hit was selected. Pulsed-field gel electrophoresis (PFGE) (5) was performed to predict the number and approximate sizes of chromosomes. The program for PFGE was 48 h at 3 V/cm with a 500-s switch time at an included angle of 106° with 0.5× Tris-borate-EDTA (TBE) on 0.75% LMP (low melting point) agarose at 14°C.

The total length of the sequenced genome was 12,238,196 bp (without N), with a GC content of 29.82%. A total of 1970.8 Mb of raw data was sequenced, representing around 145-fold coverage of the *P. tannophilus* genome. Five thousand three hundred forty-six protein-coding genes (coding sequences [CDSs]) were predicted, and 4,463 (83.5%) genes were annotated with function.

Four coding sequences of *P. tannophilus* were retrieved from GenBank to compare with the annotated CDSs. There were 2-bp differences among 4,803 bp of length. Furthermore, 13 full-length coding sequences were PCR amplified and resequenced to estimate accuracy. The total resequenced length was 21,111 bp, with 100% similarity obtained.

Based on PFGE results, six chromosomal bands were separated, with two of the bands probably migrating as doublets. The sizes of the chromosomal bands were estimated to be 2.9 ± 0.05 Mb (mean ± standard deviation), 2.1 ± 0.04 Mb, 1.9 ± 0.05 Mb, 1.6 ± 0.08 Mb (doublet), 1.3 ± 0.07 Mb (doublet), and 0.98 ± 0.02 Mb based on comparison with the yeast marker *Hansenula wingei*. The estimated genome size of *P. tannophilus* was approximately 13.6 ± 0.4 Mb with an estimated 8 chromosomes.

Nucleotide sequence accession numbers. The draft genome sequences of *P. tannophilus* were deposited in EMBL with contig accession numbers CAHV01000001 to CAHV01000267.

ACKNOWLEDGMENTS

This work was funded by the European Community's 7th Framework Research Programme under grant agreement number 213506 (project GLYFINERY), providing financial support to X.L. and M.W.

REFERENCES

1. Fu N, Peiris P. 2008. Co-fermentation of a mixture of glucose and xylose to ethanol by *Zymomonas mobilis* and *Pachysolen tannophilus*. *World J. Microbiol. Biotechnol.* 24:1091–1097.
2. Kurtzman C. 1983. Biology and physiology of the D-xylose fermenting yeast *Pachysolen tannophilus*, p 73–83. In Feichter A, Jeffries TW (ed), *Advances in biochemical engineering and biotechnology*, vol 27, Springer, Berlin, Germany.
3. Li R, et al. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20:265–272.
4. Liu X, Jensen PR, Workman M. 2012. Bioconversion of crude glycerol feedstocks into ethanol by *Pachysolen tannophilus*. *Bioresour. Technol.* 104:579–586.
5. Maringe L, Lydall D. 2006. Pulsed-field gel electrophoresis of budding yeast chromosomes, p 65–73. In Xiao W (ed), *Yeast protocol*, vol. 313, Humana Press, Totowa, NJ.
6. Sathesh-Prabu C, Murugesan AG. 2011. Potential utilization of sorghum field waste for fuel ethanol production employing *Pachysolen tannophilus* and *Saccharomyces cerevisiae*. *Bioresour. Technol.* 102:2788–2792.
7. Schneider H, Wang PY, Chan YK, Maleszka R. 1981. Conversion of D-xylose into ethanol by the yeast *Pachysolen tannophilus*. *Biotechnol. Lett.* 3:89–92.
8. Slininger PJ, Bolen PL, Kurtzman CP. 1987. *Pachysolen tannophilus*: Properties and process considerations for ethanol production from d-xylose. *Enzyme Microb. Technol.* 9:5–15.
9. Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res.* 32(Suppl 2):W309–W312.
10. Zhao L, Yu J, Zhang X, Tan T. 2010. The ethanol tolerance of *Pachysolen tannophilus* in fermentation on xylose. *Appl. Biochem. Biotechnol.* 160:378–385.

Received 13 April 2012 Accepted 16 April 2012

Address correspondence to Mhairi Workman, mwo@bio.dtu.dk.

*Present address: Mhairi Workman, Center for Microbial Biotechnology, Department of Systems Biology, Søtofts Plads, Technical University of Denmark, Kongens Lyngby, Denmark.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/EC.00114-12